

# Towards standard methods for benchmark quality ab initio thermochemistry — W1 and W2 theory

Jan M.L. Martin\* and Glênisson de Oliveira

*Department of Organic Chemistry, Kimmelman Building, Room 262, Weizmann Institute of Science, 76100 Rehovot, Israel. Email: comartin@wicc.weizmann.ac.il*

(*J. Chem. Phys.* MS# A9.02.198; received February 22, 1999; revised February 2, 2008)

## Abstract

Two new schemes for computing molecular total atomization energies (TAEs) and/or heats of formation ( $\Delta H_f^\circ$ ) of first-and second-row compounds to very high accuracy are presented. The more affordable scheme, W1 (Weizmann-1) theory, yields a mean absolute error of 0.30 kcal/mol and includes only a single, molecule-independent, empirical parameter. It requires CCSD (coupled cluster with all single and double substitutions) calculations in *spdf* and *spdfg* basis sets, while CCSD(T) [i.e. CCSD with a quasiperturbative treatment of connected triple excitations] calculations are only required in *spd* and *spdf* basis sets. On workstation computers and using conventional coupled cluster algorithms, systems as large as benzene can be treated, while larger systems are feasible using direct coupled cluster methods. The more rigorous scheme, W2 (Weizmann-2) theory, contains no empirical parameters at all and yields a mean absolute error of 0.23 kcal/mol, which is lowered to 0.18 kcal/mol for molecules dominated by dynamical correlation. It involves CCSD calculations in *spdfg* and *spdfgh* basis sets and CCSD(T) calculations in *spdf* and *spdfg* basis sets. On workstation computers, molecules with up to three heavy atoms can be treated using conventional coupled cluster algorithms,

while larger systems can still be treated using a direct CCSD code. Both schemes include corrections for scalar relativistic effects, which are found to be vital for accurate results on second-row compounds.

## I. INTRODUCTION

Thermochemical data such as molecular heats of formation are among the most crucial quantitative chemical data. Thanks to great progress made in recent years in both methodology and computer technology, a broad range of empirical, semiempirical, density functional, and ab initio schemes now exist for this purpose (for a recent collection of reviews, see Ref. [1]).

At present, only ab initio-based methods can claim ‘chemical accuracy’ ( $\pm 1$  kcal/mol). The most popular such schemes are undoubtedly the G2 [2] and G3 [3] theories of Pople and coworkers (which are based on a combination of additivity approximations and empirical corrections applied to relatively low-level calculations), followed by the CBS-Q [4–6] and CBS/APNO [4] methods which are intricate combinations of extrapolation and empirical correction schemes. With the exception of CBS/APNO (which allows for 0.5 kcal/mol accuracy, on average [7], but is restricted to first-row compounds), all these schemes allow for mean absolute errors of about 1 kcal/mol, although errors for some individual molecules (e.g.  $\text{SO}_2$ ,  $\text{SiF}_4$  [3]) can be *much* larger (e.g. about 8–12 kcal/mol for  $\text{SiF}_4$  using G2 theory, and 4 kcal/mol using G3 theory [8]).

In fact, many of the experimental data in the “enlarged G2 set” [9] employed in the parametrization of several of these methods (notably G3 theory and several of the more recent density functional methods [10]) *themselves* carry experimental uncertainties larger than 1 kcal/mol.

The aim pursued in the present work is a more ambitious one than chemical accuracy. In light of the prevalent use of kJ/mol units in the leading thermochemical tables compendia (JANAF [11] and CODATA [12]), we shall arbitrarily define a mean absolute error of one

such unit, i.e. 0.24 kcal/mol, as ‘calibration accuracy’ — with the additional constraint that no individual error be larger than the ‘chemical accuracy’ goal of 1 kcal/mol.

One of us [13,14] has recently shown that this goal is achievable for small polyatomics using present technology. The approach followed employed explicit treatment of inner-shell correlation [15], coupled cluster calculations in augmented basis sets of *spdfg* and *spdfgh* quality, and extrapolation of the valence correlation contribution to the atomization energy using formulas [13] based on the known asymptotic convergence behavior [16–18] of pair correlation energies. In this manner, total atomization energies ( $\text{TAE}_e$ ) of about 15 first-row diatomics and polyatomics for which experimental data are known to about 0.1 kcal/mol could be determined to within 0.25 kcal/mol on average *without any empirical parameters*. (Upon introducing an empirical correction for A–N bonds, this could be improved to 0.13 kcal/mol, clearly within the target.) In fact, using this method, an experimental controversy concerning the heat of formation of gaseous boron — a quantity that enters any ab initio or semiempirical calculation of the heat of formation of any boron compound — could be resolved [19] by a benchmark calculation of the total atomization energy of  $\text{BF}_3$ .

Benchmark studies along similar lines by several other groups (e.g. those of Helgaker [20], Bauschlicher [21], Dunning [22]) point in the same direction. Among those, Bauschlicher [21] was the first to suggest that the inclusion of scalar relativistic corrections may in fact be essential for accurate results on second-row molecules.

High-accuracy results on second-row compounds can only be achieved in this manner — as has been shown repeatedly [23,24,21] — if high-exponent *d* and *f* functions are added to the basis set. As shown by one of us [23], these ‘inner shell polarization functions’ address an SCF-level effect which bears little relationship to inner-shell correlation, and actually dwarfs the latter in importance (contributions as high as 10 kcal/mol having been reported [23,8]).

All these approaches carry a dual disadvantage: their extravagant computational cost and their reliance on the quantum chemical expertise of the operator.

The target of the present study was to develop computational procedures that meet the

following requirements:

- they should have mean absolute errors on the order of 0.25 kcal/mol or less, and problem molecules (if any) should be readily identifiable;
- the method should be applicable to at least first-and second-row molecules;
- it should be robust enough to be applicable in a fairly ‘black-box’ fashion by a non-specialist;
- it should rely as little as possible (preferably not at all) on empirical parameters, empirical additivity corrections, or other ‘fudges’ derived from *experimental* data;
- relatedly, it should explicitly include all the physical effects that are liable to affect molecular binding energies of first-and second-row compounds, rather than rely upon absorbing them in empirical parametrization;
- last but not least, it should be sufficiently cost-effective that a molecule the size of, e.g., benzene should be treatable on a workstation computer.

In the course of this work, we will present two schemes which we shall denote W1 and W2 (for Weizmann-1 and Weizmann-2) theories. W2 theory yields about 0.2 kcal/mol (or better) accuracy for first-and second-row molecules with up to four heavy atoms, and involves no empirical parameters. W1 theory is applicable to larger systems (we shall present benzene and trans-butadiene as examples), yet still yields a mean absolute error of about 0.30 kcal/mol and includes only a single, molecule-independent, empirical parameter which moreover is derived from calculated rather than experimental results.

## II. COMPUTATIONAL DETAILS

Most electronic structure calculations reported in this work were carried out using MOLPRO 97.3 [25] and MOLPRO 98.1 [26] running on a Silicon Graphics (SGI) Octane workstation and on the SGI Origin 2000 of the Faculty of Chemistry. The full CCSDT (coupled

cluster with all connected single, double, and triple substitutions [27]) calculations were carried out using ACES II [28] running on a DEC Alpha 500/500 workstation.

SCF and valence correlation calculations were carried out using correlation consistent [29,30] polarized  $n$ -tuple zeta (cc-pVnZ, or VnZ for short) ( $n=D, T, Q, 5, 6$ ) and augmented correlation consistent [31] polarized  $n$ -tuple zeta (aug-cc-pVnZ, or AVnZ for short) ( $n=D, T, Q, 5, 6$ ) basis sets of Dunning and coworkers. The maximum angular momentum parameter  $l$ , which occurs in the extrapolation formulas for the correlation energy, is identified throughout with the  $n$  in VnZ and AVnZ. Except for the calculation of the electron affinity of hydrogen, regular VnZ basis sets were used throughout on hydrogen atoms.

Most valence correlation calculations were carried out using the CCSD (coupled cluster with all single and double substitutions [32]) and CCSD(T) (i.e. CCSD followed by a quasiperturbative estimate of the effect of connected triple excitations [33,34]) electron correlation methods. The CCSD(T) method is known [35] to be very close to an exact solution within the given one-particle basis set if the wave function is dominated by dynamical correlation.

Where possible, imperfections in the treatment of connected triple excitations were estimated by comparing with full CCSDT calculations. The effect of connected quadruple and higher excitations were estimated by small basis set FCI (full configuration interaction) calculations — which represent exact solutions with a finite basis set.

Inner-shell correlation contributions were evaluated by taking the difference between valence-only and all-electron CCSD(T) calculations in special core-correlation basis sets. For first-row compounds, both Dunning’s ACVQZ (augmented correlation consistent core-valence quadruple zeta [36]) basis set and the Martin-Taylor (MT) core correlation basis sets [37,15] were considered; for second-row compounds only the MT basis sets. The latter are generated by completely decontracting a CVnZ or ACVnZ basis set, and adding one tight  $p$  function, three high-exponent  $d$  functions, two high-exponent  $f$  functions, and (in the case of the MTv5z basis set) one high-exponent  $g$  function to the basis set. The additional

exponents were derived from the highest ones already present for the respective angular momenta, successively multiplied by 3.0. The smallest such basis set, MTvtz (based on VTZ) is also simply denoted MT.

Scalar relativistic corrections were calculated at the ACPF (averaged coupled pair functional [38]) level as expectation values of the first-order Darwin and mass-velocity terms [39,40]. An idea of the reliability of this approach is given by comparing a very recent relativistic (Douglas-Kroll [41]) coupled cluster calculation [42] of the relativistic contribution to TAE[SiH<sub>4</sub>],  $-0.67$  kcal/mol, with the identical value of  $-0.67$  kcal/mol obtained by means of the present approach. For GaCl, GaCl<sub>2</sub>, and GaCl<sub>3</sub> — where relativistic effects are an order of magnitude stronger than even in the second-row systems considered here — Bauschlicher [43] found that differences between Douglas-Kroll calculations and the presently followed approach amounted to 0.12 kcal/mol or less on the binding energy.

Spin-orbit coupling constants were evaluated at the CASSCF-CI level using the *spdf* part of the MTav5z basis set. (For a recent review of the methodology involved, see Ref. [44].)

Density functional calculations for the purposes of obtaining certain reference geometries and zero-point energies were carried out using the Gaussian 98 [45] package. Both the B3LYP (Becke 3-parameter [46]-Lee-Yang-Parr [47]) and B3PW91 (Becke 3-parameter [46]-Perdew-Wang-1991 [48]) exchange-correlation functionals were considered.

Most geometry optimizations were carried out at either the CCSD(T)/VQZ+1 or the B3LYP/VTZ+1 (in some cases B3PW91/VTZ+1) levels of theory, where the notation  $VnZ+1$  indicates the addition to all second-row atoms of a single high-exponent  $d$ -type ‘inner polarization function’ [49,23] with an exponent equal to the highest  $d$  exponent in the Dunning V5Z basis set. In the past this was found [49,23,50,51] to recover the largest part of the effects of inner polarization on geometries and vibrational frequencies. (We note that for molecules consisting of first-row atoms only, the  $VnZ+1$  basis sets are equivalent to regular  $VnZ$  basis sets.)

Past studies [52–54] of the convergence behavior of the SCF energy have shown it to be

very well described by a geometric extrapolation of the type first proposed by Feller [55],  $A + B/C^l$ . Clearly, for this purpose a succession of three SCF/AV $n$ Z basis sets is required.

For the valence correlation CCSD and (T) energies, two extrapolation formulas were considered. The first,  $A + B/(l + 1/2)^\alpha$ , was proposed by Martin [13] — the philosophy being that using the extrapolation exponent as an adjustable parameter would enable inclusion of higher-order terms in the asymptotic expansion

$$A/(L + 1)^3 + B/(L + 1)^4 + C/(L + 1)^5 + \dots \quad (1)$$

while the denominator shift of  $1/2$  was a compromise — for identification of the  $l$  in cc-pV $l$ Z with  $L$  — between hydrogen and nonhydrogen atoms. The second formula, simply  $A + B/l^3$ , was proposed by Helgaker and coworkers [56] — where  $l$  was identified with  $L - 1$  throughout. Halkier et al. [56] already noted that in terms of the extrapolated energy using  $A + B/(l + C)^D$ , the parameters  $C$  and  $D$  were very strongly coupled, and that it only made sense to vary one of them.

The combination of treatments for SCF, CCSD valence correlation, (T), imperfections in the T treatment, and connected quadruple and higher excitations is compactly denoted here by W[p5;p4;p3;p2;p1], in which p1 denotes the basis sets involved in the SCF extrapolation, p2 the basis sets involved in the CCSD extrapolation, p3 those in the (T) extrapolation (which may or may not be different from p2), p4 (if present) the basis sets used in correcting for imperfections in the treatment of connected triple excitations, and p5 (if present) those involved in evaluating the effect of connected quadruple and higher excitations. If any of the p's consists of a single index, a simple additivity approximation is implied; two indices denote a two-parameter extrapolation of the type  $A + B/l^3$ , while three indices indicate a three-parameter extrapolation of the type  $A + B/(l + 1/2)^\alpha$  in the case of correlation contributions, and  $A + B/C^l$  in the case of SCF contributions. For example, the level of theory used in the previous work of Martin and Taylor would be W[TQ5;TQ5;TQ5] in the present notation, while W[D;Q;TQ5;TQ5;TQ5] indicates W[TQ5;TQ5;TQ5]+CCSDT/AVQZ-CCSD(T)/AVQZ+FCI/AVDZ-CCSDT/AVDZ.

### III. ATOMIC ELECTRON AFFINITIES AS A ‘LITMUS TEST’

The electron affinities of the first-and second-row atoms have often been used as benchmarks for high-level electronic structure methods (see e.g. the introductions to Refs. [57,58] for reviews). Because electron affinities involve a change in the number of electrons correlated in the system, they are very taxing tests for any electron correlation method; in addition, they involve a pronounced change in the spatial extent of the wave function, making them very demanding in terms of the basis set as well.

Until recently, three of the first-and second-row atomic electron affinities were imprecisely known experimentally (B, Al, and Si): this situation was changed very recently by high-precision measurements for recent experiments for B [59], Al [60,61], and Si [62].

The approach we have chosen here for the SCF and valence correlation components is summarized in our notation as  $W[n,Q,56,56,Q56]$  for the first-row atoms, and  $W[n,Q,Q5,Q5,TQ5]$  for the second-row atoms. The effect of inner-shell correlations was assessed at the CCSD(T)/MTav5z level, while Darwin and mass-velocity corrections were evaluated at the ACPF/MTav5z level. Finally, spin-orbit splittings were calculated at the CASSCF-CI level with the *spdf* part of a MTav5z basis set. (For technical reasons, the *h* functions were omitted in both the scalar relativistic and spin-orbit calculations, as were the *g* functions in the latter.)

Our best computed results are compared with experiment in Table I, where results from recent calibration studies are also summarized.

Agreement between computed and observed values can be described without reservation as excellent: the mean absolute error amounts to 0.0009 eV. The fact that this accuracy is obtained systematically and across the board strongly suggest that the ‘right result was obtained for the right reason’. Upon eliminating the corrections for imperfections in CCSD(T), i.e. restricting ourselves to  $W[56,56,Q56]$  for first-row atoms and  $W[Q5,Q5,TQ5]$  for second-row atoms, the mean absolute error increases by an order of magnitude to 0.009 eV, i.e. about 0.2 kcal/mol. As we shall see below, this is essentially the type of accuracy we can



obtain for molecules without corrections for CCSD(T) imperfections.

The importance of Darwin and mass-velocity corrections increases, as expected, with increasing  $Z$ , and its contribution becomes quite nontrivial for atoms like Cl. It is therefore to be expected that, e.g., in polar second-row molecules like ClCN or SO<sub>2</sub> they will contribute substantially to TAE as well.

The importance of inner-shell correlation effects is actually largest for Al, because of the small gap between valence and sub-valence orbitals in the early second-row elements.

Table II compares the convergence behavior of the extrapolated valence correlation contributions as a function of the largest basis set used, both using the Martin three-term and Helgaker two-term formulas. While both formulas appear to give the same answer if the underlying basis sets are large enough, the two-term formula is by far the more stable towards reduction of the basis sets used in the extrapolation. Since the use of W[Q56,Q56,Q56] is hardly an option for molecules, the two-term formula appears to be the formula of choice.

Following the suggestion of a referee, we have considered (Table II) the performance of some other extrapolation formulas for the valence correlation energy. As a point of reference, we have taken an “experimental valence correlation contribution to EA”, which we derived by subtracting all computed contributions other than the valence correlation energy from the best experimental EAs. While some residual uncertainties may remain in some of the individual contributions, these should be reliable to 0.001 eV on average.

As seen in Table II, performance of the geometric series extrapolation [55]  $A + B/C^n$  is outright poor: in fact, for extrapolation from AV{D,T,Q}Z results the error is twice as large as that caused by not extrapolating at all. If AV{T,Q,5}Z basis sets are used, mean absolute error drops to 0.015 eV, which is still an order of magnitude larger than for the  $A + B/l^3$  extrapolation, and only slightly better than not extrapolating at all. Finally, for AV{Q,5,6}Z basis sets, the error is three times smaller than complete omission of extrapolation, but three times larger than that of using any of the following formulas:  $A + B/l^3$  [56],  $A + B/l^C$  [13], or  $A + B/(l + 1/2)^4 + C/(l + 1/2)^6$  [13]. All three of the latter yield a mean absolute

error of about 0.001 eV, on about the same order of accuracy as the reference values. For the smallest basis set series AV{D,T,Q}Z, the mixed exponential-Gaussian extrapolation [63]  $A + B/\exp(l-1) + C/\exp((l-1)^2)$  represents a very substantial improvement over  $A + B/C^l$ , and actually exhibits slightly better performance than  $A + B/l^3$ . For the AV{T,Q,5}Z series which is of greatest interest here, the Halkier et al.  $A + B/l^3$  formula by far outperforms the other formulas considered.

In short, it appears to be established that the two-term formula of Helgaker and coworkers [56] is the extrapolation method of choice overall, with the Martin three-term formulas the second-best choice provided basis sets of AV{T,Q,5}Z quality are used. The mixed exponential-Gaussian formula performs slightly better than  $A + B/l^3$  if only AV{D,T,Q}Z basis sets are used (see however Section VI.C below).

Computed spin-orbit contributions to the electron affinities are compared in Table III to values obtained from observed fine structures [64,59,61]. While small deviations appear to persist, these may at least in part be due to higher-order spin-orbit effects which were neglected in the calculation rather than to deficiencies in the electronic structure treatment. At any rate, to the accuracy relevant for our purpose (establishing spin-orbit corrections to molecular binding energies) it appears to be immaterial whether the computed or the experimentally derived values are used.

Finally, the convergence of the SCF component is so rapid that it appears to be essentially irrelevant which extrapolation formula is used — the amount bridged by the extrapolation is on the order of 0.0001 eV.

#### IV. RESULTS FOR MOLECULES

Since application of electron correlation methods more elaborate than CCSD(T) would be well-nigh impossible for molecules of practical size, we have restricted ourselves to W[Q5;Q5;TQ5] and W[TQ5;TQ5;TQ5].

Inner-shell correlation contribution, as well as scalar relativistic corrections, were initially

computed with the largest basis sets practicable – in most cases ACV5Z or MTavqz (see Table IV for details).

From a prerelease version of a re-evaluation of the experimental data in the G2/G3 set at the National Institute for Standards and Technology (NIST) [65], we have selected 28 first-and second-row molecules which satisfy the following criteria: (a) the uncertainty in the experimental total atomization energy TAE is on the order of 0.25 kcal/mol or better; (b) the molecules are not known to exhibit severe multireference effects; (c) anharmonic vibrational zero-point energies are available from either experiment or high-level ab initio calculations (see footnotes to Table IV for details).

Geometries were optimized at the CCSD(T)/VQZ+1 level, and to all second-row atoms a complement of two tight d and one tight f function were added in every basis set to ensure saturation in inner-shell polarization effects. In all cases, the exponents were derived as even-tempered series  $\alpha\beta^n$  with  $\beta = 3.0$  and  $\alpha$  the highest exponent already present for that angular momentum.

Computed (W[Q5;Q5;TQ5]) and observed results are compared in Table IV. The excellent agreement between theory and experiment is immediately apparent: in many cases, the computed results fall within the already quite narrow experimental error bars. Over the entire sample of molecules, the mean absolute error is 0.24 kcal/mol, with the largest errors being about 0.6 kcal/mol ( $\text{O}_2$  and  $\text{F}_2$ ). Restricting our sample to first-row molecules only, we find a mean absolute error of 0.24 kcal/mol, which however gets reduced to 0.17 kcal/mol (maximum error 0.39 kcal/mol for  $\text{N}_2$ ) upon elimination of  $\text{F}_2$ ,  $\text{NO}$ , and  $\text{O}_2$  as having known appreciable nondynamical correlation effects. Over the subset of second-row molecules in our sample MAE is 0.23 kcal/mol (maximum error 0.44 kcal/mol for  $\text{H}_2\text{S}$ ); upon elimination of  $\text{H}_2\text{S}$  and  $\text{SO}_2$  this is lowered to 0.20 kcal/mol.

It should be noted that these MAEs are comparable to those found by Martin and Taylor [14] for a sample of first-row molecules, yet unlike their study no correction for N-containing bonds is required here.

The possibility that the errors in  $F_2$ ,  $NO$ , and  $O_2$  are actually due to residual basis set incompleteness and/or that the excellent agreement with experiment for the other molecules is actually due to an error compensation involving deficiencies in the predicted basis set limit, was examined by carrying out  $W[56;56;Q56]$  calculations for  $H_2O$ ,  $F_2$ ,  $NO$ ,  $O_2$ ,  $N_2$ ,  $HF$ , and  $CO$ . As seen in Table V, the predicted basis set limits do not differ materially from their  $W[Q5;Q5;TQ5]$  counterparts, strongly suggesting that the latter expression in fact *does* reach the basis set limit and that the residual errors are largely due to imperfections in the CCSD(T) method.

While molecules liable to exhibit such errors are readily identifiable from inspection of the largest coupled cluster amplitudes or evaluation of the  $\mathcal{T}_1$  diagnostic [66], an even simpler criterion is apparently offered by the ratio  $TAE[SCF]/TAE[SCF+val.corr.]$ . In “well-behaved” molecules such as  $CH_4$  and  $H_2O$ , the SCF component makes up upwards of two-thirds of the binding energy, while in  $NO$  and in  $O_2$  it makes up no more than a third and a fifth, respectively, of the total and  $F_2$  is actually metastable towards dissociation at the SCF level. While for some molecules of this variety we actually obtain excellent results (e.g.  $ClF$ ), this may be due to error compensation or to the binding energies being fairly small to begin with.

Further inspection of Table IV reveals that some of the ‘negligible’ contributions are in fact quite significant at the present precision level: for instance, Darwin and mass-velocity contributions in  $SO_2$  amount to -0.71 kcal/mol (for  $SiF_4$  a somewhat extravagant -1.88 kcal/mol was found [8]), while atomic spin-orbit splitting in such compounds as  $Cl_2$ ,  $ClF$ , and  $SO_2$  amounts to -1.68, -1.23, and -1.01 kcal/mol, respectively. Inner-shell correlation contributions of 2.36 ( $C_2H_4$ ), 2.44 ( $C_2H_2$ ), 1.68 ( $OCS$ ), and 1.76 ( $ClCN$ ) kcal/mol speak for themselves; interestingly (as noted previously [23,50]), these effects on the whole do not seem to be more important in second-row than in first-row compounds.

Finally, we shall compare the performance of  $W[TQ5;TQ5;TQ5]$  and  $W[Q5;Q5;TQ5]$  (Table VI). In general, the results with the three-point valence correlation extrapolation are

at best of the same quality as those with the two-point valence correlation extrapolation and in many cases agree less well with experiment. We therefor will use the two-point extrapolation exclusively henceforth.

## V. W2 THEORY AND ITS PERFORMANCE

Having established that our 'base level of theory' can obtain the right results for the right reason, we shall now proceed to consider simplifications.

### A. Inner-shell correlation

The use of the smaller MT basis set for the scalar relativistic contributions is found to have an effect of about 0.01 kcal/mol or less, with 0.02 kcal/mol being the largest individual cases. This approximation can therefore safely be made.

Using the same MT basis set for the core correlation contribution on average affects energetics by 0.03 kcal/mol, the largest individual effects being 0.07 kcal/mol for H<sub>2</sub>S, and 0.08 kcal/mol for OCS.

Even so, in fact, the core correlation calculations are quite CPU-time consuming, particularly for second-row compounds, due to the large number of electrons being correlated. Any further reduction would obviously be welcome — it should be noted that the MT basis set was developed not with efficiency, but with saturation (in the core-valence correlation energy) in mind. Further experimentation revealed that the tightest p, d, and f functions could safely be eliminated, but that further basis set reductions adversely affect the quality of the core correlation contribution computed. The reduced basis set shall be denoted as MTsmall, and in fact consists of a completely decontracted cc-pVTZ basis set with two tight *d* and one tight *f* functions added. Since this basis set only has about half the basis functions of the ACVQZ basis set per heavy atom, it represents a very significant CPU time savings (about 16 times) in a CCSD(T) calculation. The only molecule for which we see a

substantial difference with the MT basis set is  $\text{SO}_2$ , for which Bauschlicher and Ricca [21] previously noted that the inner-shell correlation contribution is unusually sensitive to the basis set.

For the evaluation of the Darwin and mass-velocity corrections, differences with the larger MT basis set are less than 0.01 kcal/mol across the board.

A further reduction in CPU time for the core correlation contribution would have been achieved if MP2 or even CCSD calculations could be substituted for their CCSD(T) counterparts. However, as seen from Table VII, CCSD underestimates the CCSD(T) core correlation contributions for several molecules by as much as 50%. The behavior of MP2 is quite similar and the MP2–CCSD differences are substantially smaller than the (T) contribution, suggesting that it is the treatment of connected higher excitations that is the issue. Predictably, the largest (T) effects on the core correlation contribution occur in molecules where connected triple excitations are important for the *valence* binding energy as well, e.g.  $\text{SO}_2$ ,  $\text{F}_2$ ,  $\text{Cl}_2$ ,  $\text{N}_2$ . Conversely, in  $\text{CH}_3$  or  $\text{CH}_4$ , which have quite small (T) contributions to the binding energy, CCSD does perform excellently for the core correlation contribution. In  $\text{PH}_3$  and  $\text{H}_2\text{S}$ , on the other hand, substantial errors in the core correlation are seen even as the (T) contribution to the valence correlation binding energy is quite small — it should be noted, however, that both the absolute inner-shell correlation energy and the (T) contribution to it are much more important in these second-row systems than in their first-row counterparts.

One may rightly wonder whether the inner-shell contributions are in fact converged at the CCSD(T) level. Unfortunately, if a more elaborate treatment is already impractical for the valence correlation, this would *a fortiori* be true for the inner-shell correlation. We did, however, carry out a CCSDT/MTsmall calculation on the  $\text{N}_2$  molecule, which we chose as a representative case of failure of the CCSD approach for core correlation. The resulting CCSDT level core contribution, 0.87 kcal/mol, is only 0.05 kcal/mol larger than the CCSD(T) value of 0.82 kcal/mol, to be compared with 0.42 kcal/mol at the MP2 and 0.52 kcal/mol at the CCSD level. It cannot be ruled out a priori that connected quadruple and

higher excitations might contribute to the inner-shell correlation energy. However, since apparently their importance for the valence correlation binding energy is very small (otherwise a treatment that completely ignores them would not yield the type of agreement with experiment found in this work), it seems unlikely that they would greatly contribute to the inner-shell correlation energy.

The “G3large” basis set used to evaluate, among other things, inner-shell correlation effects in G3 theory [3] is still smaller than the MTsmall basis set, and its performance therefore is certainly of interest. Alas, in Table VII it is seen that in many cases it seriously overestimates the inner-shell correlation energy, almost certainly because of basis set superposition error which [21] is apparently more of an issue for inner-shell correlation energies than for their valence counterparts. In G3 theory, the inner-shell correlation is evaluated at the MP2 level: hence the two errors cancel to a substantial extent.

### B. Zero-point energy

Not in all cases is a complete anharmonic force field calculation feasible. We find in Table VIII that B3LYP/cc-pVTZ+1 *harmonic* zero-point energies scaled by 0.985 reproduce the rigorous *anharmonic* zero-point energies quite nicely. (The scaling factor is about halfway between what would be required for fundamentals, about 0.97, and harmonics, about 1.00. [67,68])

### C. Separate extrapolation of CCSD and (T)

The (T) contribution makes up a relatively small part of the valence correlation energy, while its evaluation, in large basis sets and for systems with very many electrons, will dominate the CPU time. For instance, in a very recent study on SiF<sub>4</sub> [8], a CCSD(T) calculation with an AVQZ basis set on F and a VQZ+2d1f basis set on Si took 50h7' using MOLPRO on an SGI Octane workstation (768 MB of memory being allocated to the job), of which

41h30' were spent in the (T) step alone.

In addition, it was previously noted by Helgaker and coworkers [69] that the (T) contribution appears to converge faster with the basis set than the CCSD correlation energy, for which reason they actually propose its separate evaluation in a smaller basis set. In the present case, we have considered extrapolating it from AVTZ+2d1f and AVQZ+2d1f results rather than AVQZ+2d1f and AV5Z+2d1f, respectively. In our adopted notation, this becomes  $W[\text{TQ};\text{Q5};\text{TQ5}]$ .

As seen in Table VI, the difference in quality between  $W[\text{TQ};\text{Q5};\text{TQ5}]$  and  $W[\text{Q5};\text{Q5};\text{TQ5}]$  appears to be essentially negligible. This is an important conclusion, since it means that the largest basis set calculation to be carried out is only at the CCSD level, at a fraction of the cost of the full CCSD(T) counterpart — moreover it can be done using direct algorithms. [70]

#### D. Protocol for W2 theory

The protocol obtained by introducing the succesful approximations given above will be denoted here as W2 (Weizmann-2) theory. Its steps consist of the following:

- geometry optimization at the CCSD(T)/VQZ+1 level, i.e. CCSD(T)/VQZ if only first-row atoms are present;
- zero-point energy obtained from a CCSD(T)/VTZ+1 anharmonic force field or, failing that, B3LYP/VTZ+1 frequencies scaled by 0.985 (vide infra);
- Carry out CCSD(T)/AVTZ+2d1f and CCSD(T)/AVQZ+2d1f single-point calculations;
- Carry out a CCSD/AV5Z+2d1f single-point calculation;
- the SCF component of TAE is extrapolated by  $A + B/C^l$  from SCF/AVTZ+2d1f, SCF/AVQZ+2d1f, and SCF/AV5Z+2d1f results ( $l=3, 4$ , and  $5$ , respectively);



- the CCSD valence correlation component is obtained from applying  $A + B/l^3$  to CCSD/AVQZ+2d1f; and CCSD/AV5Z+2d1f valence correlation energies ( $l=4$  and  $5$ , respectively). It is immaterial whether this is done on total energies or on components to TAE;
- the (T) valence correlation component is obtained from applying  $A + B/l^3$  results to CCSD(T)/AVTZ+2d1f and CCSD(T)/AVQZ+2d1f values for the (T) contribution. It is again immaterial whether this is done on total energies or on components to TAE;
- core correlation computed at CCSD(T)/MTsmall level
- scalar relativistic corrections (and, if necessary, spin-orbit splittings) computed at ACPF/MT level. To save CPU time, this can be combined into a single job with the previous step.

On a typical workstation at the time of writing (e.g. an SGI Octane with 1 GB of RAM and  $2 \times 18$  GB external disks) its applicability range would be about three heavy atoms in  $C_{2v}$  symmetry, although the main limiting factor would be disk space and larger systems could be treated if a direct CCSD code were available.

We will illustrate the CPU time savings made in the W2 approach, compared to our most rigorous calculations, using two examples: a first-row diatomic (CO) and a second-row molecule (OCS). Using MOLPRO on an SGI Octane workstation, the most accurate calculations reported in this work (Table IV) required 21h36' for CO and no less than 362h12' for OCS. W2 theory yields essentially identical results at a cost of 1h12' (CO) or 13h42' (OCS) — a reduction by a factor of 20–30 which is typical of the other molecules.

## VI. W1 THEORY AND ITS PERFORMANCE

In an effort to obtain a method that is applicable to larger systems, we introduce a few further approximations. Relevant results can be found in Table VI.

### A. Use of density functional reference geometries

B3LYP/cc-pVTZ+1 geometries are close enough to their CCSD(T)/VQZ+1 counterparts that their use does not cause a major effect on the final computed result. There is one notable exception to this rule for the molecules considered here:  $\text{Cl}_2$ , for which the B3LYP/VTZ+1 bond distance of 2.0130 Å is quite different from its CCSD(T)/VQZ+1 counterpart, 1.9972 Å. (The experimental value [71] is 1.987<sub>9</sub> Å.) While B3LYP and B3PW91 on the whole tend to produce essentially identical geometries and harmonic frequencies [68], it has been argued previously [72] that the B3PW91 functional may be somewhat more reliable for systems with high electron density; and in fact, the B3PW91/VTZ+1 bond distance of 1.9912 Å is much closer to the CCSD(T)/VQZ+1 and experimental value.

Even so, the use of a B3LYP/VTZ+1 reference geometry still does not affect the computed  $D_e$  by more than 0.14 kcal/mol. We conclude that CCSD(T) geometry optimizations, which will become fairly costly for larger molecules, can safely be replaced by B3LYP calculations, which can also serve for obtaining zero-point energies.

### B. Further reduction of basis set sizes

The obvious first suggestion would be to also carry out the CCSD extrapolation using smaller basis sets, i.e. W[TQ,TQ,DTQ].

The effect for the SCF component of TAE is very small on condition that the extrapolation is carried out *not* on the individual total energies but rather on the computed SCF components of TAE themselves. Clearly error compensation occurs between the molecule and the constituent atoms.

The effect on the valence correlation component, unfortunately, is rather more significant. Over the ‘training set’, MAE rises to 0.37 kcal/mol even after  $\text{SO}_2$  (which clearly is a pathological case here) has been eliminated. Aside from the latter, eliminating systems with mild nondynamical correlation does not lead to any significant reduction in MAE. Also noteworthy

is that, on average, the binding energies appear to be somewhat overestimated: this is easily explained from the fact that basis sets like AVTZ+2d1f are not quite saturated in the *radial* correlation energy either, and that therefore the  $\text{TAE}[\text{AVQZ}+2\text{d1f}] - \text{TAE}[\text{AVTZ}+2\text{d1f}]$  gap will be an overestimate of the  $\text{TAE}[\text{L}=4] - \text{TAE}[\text{L}=3]$  gap.

### C. Use of empirical extrapolation exponents

Truhlar [73] considered the use of  $L$ -extrapolation formulas with *empirical* exponents, carried out from cc-pVDZ and cc-pVTZ calculations, as an inexpensive alternative to very large basis set calculations.

We will investigate here a variant of this suggestion adapted to the present framework. The valence correlation component to TAE will indeed be extrapolated using the formula  $A + B/l^\beta$ , in which  $\beta$  is now an empirical parameter — we will denote this  $\text{W}[\text{Q5};\text{Q5};\text{TQ5}]\beta$  and the like.

We then add in all the further corrections (core correlation, scalar relativistics, spin-orbit) that occur in W2 theory, and try to determine  $\beta$  by minimizing MAE with respect to the experimental TAE values for our ‘training set’. Not surprisingly, for  $\text{W}[\text{Q5};\text{Q5};\text{TQ5}]\beta$  this yields an optimum exponent ( $\beta=2.98$ ) which differs insignificantly from the ‘ideal’ value of 3.0. Alternatively,  $\beta$  could be optimized for the best possible overlap with the  $\text{W}[56;56;\text{Q56}]$  results: in fact, the same conclusion is obtained, namely that making  $\beta$  an empirical parameter does *not* improve the quality of the results.

For  $\text{W}[\text{TQ};\text{TQ};\text{DTQ}]$  however, the situation is rather different. The optimum exponent  $\beta$  is found to be 3.18 if optimized against the experimental TAE values, and 3.16 (insignificantly different) if optimized against the  $\text{W}[\text{Q5};\text{Q5};\text{TQ5}]$  results. In both cases, MAE drops to 0.30 kcal/mol, and on average no more overestimation occurs.

$\text{W}[\text{TQ};\text{TQ};\text{DTQ}]\beta$  represents a significant savings over W2 theory. Its time-determining step in molecules with many electrons will be the evaluation of the parenthetical triples in the AVQZ+2d1f basis set — their elimination would be most desirable.

A natural suggestion would then be  $W[DT;TQ;DTQ]\beta$ . Optimization of  $\beta$  against the experimental TAE values yields  $\beta=3.26$ ; the not greatly different  $\beta=3.22$  is obtained by minimization of the deviation from  $W[Q5;Q5;TQ5]$  results for the training set. Since the latter does not explicitly depend on experimental results and therefore minor changes in the computational protocol do not require recalculation for the entire ‘training set’, we will opt for the latter alternative.

In either case, we obtain MAE=0.30 kcal/mol — for a calculation that requires not more than an AVTZ+2d1f basis set for the largest CCSD(T) calculation, and an AVQZ+2d1f basis set for the largest CCSD calculation. Again, the latter is amenable to a direct algorithm.

#### D. Protocol for W1 theory

We thus propose the following protocol for a computational level which we will call W1 (Weizmann-1) theory:

- geometry optimization at the B3LYP/VTZ+1 level (B3LYP/VTZ if only first-row atoms are present). Alternatively, the B3PW91 exchange-correlation functional may be preferable for some systems like  $\text{Cl}_2$  — under normal circumstances, B3LYP/VTZ+1 and B3PW91/VTZ+1 should yield virtually identical geometries;
- zero-point energy obtained from B3LYP/VTZ+1 (or B3PW91/VTZ+1) harmonic frequencies scaled by 0.985;
- Carry out CCSD(T)/AVDZ+2d and CCSD(T)/AVTZ+2d1f single-point calculations;
- Carry out a CCSD/AVQZ+2d1f single-point calculation;
- the SCF component of TAE is extrapolated by  $A + B/C^l$  from SCF/AVDZ+2d, SCF/AVTZ+2d1f, and SCF/AVQZ+2d1f components of TAE ( $l=2, 3$ , and  $4$ , respectively)
- set  $\beta=3.22$

- the CCSD valence correlation component is obtained from applying  $A + B/l^\beta$  to CCSD/AVTZ+2d1f and CCSD/AVQZ+2d1f valence correlation energies ( $l=3$  and  $4$ , respectively). In both this and the next step, it is immaterial whether the extrapolation is carried out on components to the total energy or to TAE;
- the (T) valence correlation component is obtained from applying  $A + B/l^\beta$  results to CCSD(T)/AVDZ+2d and CCSD(T)/AVTZ+2d1f values for the (T) contribution.
- core correlation contributions are obtained at the CCSD(T)/MTsmall level;
- scalar relativistic and, where necessary, spin-orbit coupling effects are treated at the ACPF/MTsmall level. As in W2 theory, this latter step can be combined in a single job with the previous step.

W1 theory can be applied to fairly large systems (see below). CPU times are dominated by the inner-shell correlation contribution (particularly for second-row compounds), which is reflected in the relatively small time reduction compared to W2 theory — e.g., from 1h12' to 24' for CO and from 13h42' to 8h48' for OCS. In addition — contrary to W2 theory — W1 theory exhibits a pronounced difference in performance between first-row and second-row compounds: for the species in Table VI, MAE is 0.26 kcal/mol for first-row, but 0.40 kcal/mol for second-row compounds. Since the CPU time gap between W1 and W2 theory is fairly narrow for second-row species, we conclude that for accurate work on second-row species — unless precluded by disk space or memory limitations — it may well be worth to ‘walk the extra mile’ and carry out a W2 rather than a W1 calculation. For first-row systems, on the contrary, W1 may well seem the more attractive of the two.

## VII. SAMPLE APPLICATIONS TO LARGER SYSTEMS

By way of illustration, we have carried out some W1 theory calculations on trans-1,3-butadiene and benzene. All relevant computed and observed results are summarized in

Table IX. The example of benzene is representative and will be discussed here in detail — it should be mentioned that the calculation was carried out in its entirety on an SGI Octane workstation with 2x18 GB external SCSI disks.

The reference geometry was obtained at the B3LYP/cc-pVTZ level [74]. The zero-point energy at that level, after scaling by 0.985, is found to be 62.04 kcal/mol.

The SCF component of TAE is predicted to be 1044.95 kcal/mol at the one-particle basis set limit, of which only 0.39 kcal/mol is covered by the geometric extrapolation. Of the CCSD valence correlation component of 291.07 kcal/mol, however, some 10.11 kcal/mol is covered by the extrapolation, which also accounts for 2.13 kcal/mol out of the 26.55 kcal/mol connected triple excitations contribution.

The inner-shell correlation contribution is quite sizable at 7.09 kcal/mol, although this number is not qualitatively different from that for three acetylenes or three ethylenes. Finally, Darwin and mass-velocity terms contribute a small but significant -0.96 kcal/mol, and atomic spin-orbit splitting another -0.51 kcal/mol. All adds up to 1367.95 kcal/mol at the bottom of the well, or 1305.92 kcal/mol at 0 K, which is in excellent agreement with the experimental value of  $1306.1 \pm 0.12$  kcal/mol from the NIST WebBook [75].

The CCSD/VQZ calculation took 10h10' with MOLPRO on the Octane, the CCSD(T)/VTZ calculation 1h48' on a single CPU on the Origin 2000. By far the most time-consuming part of the calculation was the inner-shell correlation contribution, at 67h46', to which another 4h52' should be added for the Darwin and mass-velocity contribution. We see similar trends in the results for trans-butadiene, which agree with experiment to virtually within the stated experimental uncertainty; for allene, we obtain a value intermediate between the two experimental values proposed in the WebBook.

We find for both molecules that the sum of core-correlation and relativistic contributions can be quite well estimated by additivity approximations. For instance, the core correlation and scalar relativistic contributions with the same basis set for  $C_2H_4$  are +2.360 and -0.330 kcal/mol, respectively, adding up to 2.030 kcal/mol. Assuming 2 and 3 times this 'C=C

bond equivalent’ for butadiene and benzene, respectively, yields estimated contributions of 4.06 (trans-butadiene), and 6.09 (benzene) kcal/mol, which agree excellently with the directly computed values of 4.02 and 6.12 kcal/mol, respectively. Considering that inner-shell correlation effects should be fairly local in character, such schemes should work quite well for larger organic systems where the valence calculation would still be feasible but the explicit inner-shell calculation would not be.

## VIII. CONCLUSIONS

We have developed and presented two quasi-‘black box’ schemes for high-accuracy calculation of molecular atomization energies or, equivalently, molecular heats of formation, of first-and second-row compounds.

The less expensive scheme, W1 (Weizmann-1) theory, yields a mean absolute error of 0.30 kcal/mol and includes only a single, molecule-independent, empirical parameter. It requires no larger-scale calculations than CCSD/AVQZ+2d1f and CCSD(T)/AVTZ+2d1f (or, for nonpolar first-row compounds, CCSD/VQZ and CCSD(T)/VTZ). On workstation computers and using conventional coupled cluster algorithms, systems as large as benzene can be treated, while larger systems are feasible using direct coupled cluster methods.

The more expensive scheme, W2 (Weizmann-2) theory, contains no empirical parameters at all and yields a mean absolute error of 0.23 kcal/mol, which is lowered to 0.18 kcal/mol for molecules dominated by dynamical correlation. On workstation computers, molecules with up to three heavy atoms can be treated using conventional coupled cluster algorithms, while larger systems can still be treated using a direct CCSD code.

The inclusion of scalar relativistic (Darwin and mass-velocity) corrections is essential for good results in second-row compounds, particularly highly polar ones. Inclusion of inner-shell correlation contributions is absolutely essential: the basis set denoted as MTsmall (for Martin-Taylor small) appears to represent the best compromise between quality and computational expense. We do not recommend the use of lower-level electron correlation

methods than CCSD(T) for the evaluation of the inner-shell contribution.

Among the several infinite-basis set extrapolation formulas for the correlation energy examined, the three-parameter  $A + B/(l + 1/2)^\alpha$  expression proposed by Martin [13] and the  $A + B/l^3$  expression proposed by Helgaker and coworkers [56] yield the best results for sufficiently large basis sets, with the latter formula to be preferred on grounds of stability of the extrapolated results with the basis sets used. Geometric and mixed geometric-Gaussian extrapolation formulas [55,63] are unsatisfactory when applied to the correlation energy, although they appear to be appropriate for the SCF component.

The main limiting factor for the quality of our calculations at this stage appears to be imperfections in the CCSD(T) method. This assertion is supported by the fact that the mean absolute error in the computed electron affinities of the atoms H, B-F and Al-Cl drops from 0.009 eV to 0.0009 eV if CCSDT and full CI corrections are included.

Extrapolation of the (T) contribution to the correlation energy can, at no loss in accuracy, be carried out using smaller basis sets than the CCSD contribution.

## ACKNOWLEDGMENTS

JM is a Yigal Allon Fellow, the incumbent of the Helen and Milton A. Kimmelman Career Development Chair (Weizmann Institute), and an Honorary Research Associate (“Onderzoeksleider in eremandaat”) of the National Science Foundation of Belgium (NFWO/FNRS). GdO acknowledges the Feinberg Graduate School (Weizmann Institute) for a Postdoctoral Fellowship. This research was supported by the Minerva Foundation, Munich, Germany. The authors acknowledge enlightening discussions with (in alphabetical order) Drs. C. W. Bauschlicher Jr. (NASA Ames Research Center), Dr. Thom H. Dunning (PNNL), Prof. Trygve Helgaker (U. of Oslo, Norway), Dr. Frank Jensen (Odense U., Denmark), Dr. Timothy J. Lee (NASA Ames Research Center), and Prof. Peter R. Taylor (UCSD and National Partnership for Advanced Computing Infrastructure), and thank Dr. Peter Stern for technical assistance with various computer systems.



## REFERENCES

- [1] *Computational Thermochemistry* (Irikura, K. K. and Frurip, D. J., Eds.), ACS Symposium Series, Nr. 677, American Chemical Society, Washington, DC, 1998.
- [2] L. A. Curtiss, K. Raghavachari, G. W. Trucks, and J. A. Pople, *J. Chem. Phys.* **94**, 7221 (1991)
- [3] L. A. Curtiss, K. Raghavachari, P. C. Redfern, V. Rassolov, and J. A. Pople, *J. Chem. Phys.* **109**, 7764 (1998)
- [4] J. A. Montgomery Jr., J. W. Ochterski, and G. A. Petersson, *J. Chem. Phys.* **101**, 5900 (1994) and references therein
- [5] J. W. Ochterski, G. A. Petersson, and J. A. Montgomery Jr., *J. Chem. Phys.* **104**, 2598 (1996)
- [6] J. A. Montgomery, Jr., M. J. Frisch, J. W. Ochterski, and G. A. Petersson, *J. Chem. Phys.* **110**, 2822 (1999)
- [7] J. M. L. Martin, *Benchmark studies on small molecules*, in *Encyclopedia of Computational Chemistry*, 5 vols. (Wiley, NY, 1998)
- [8] J. M. L. Martin and P. R. Taylor, *J. Phys. Chem. A* **103**, 4427 (1999).
- [9] L. A. Curtiss, K. Raghavachari, P. C. Redfern, and J. A. Pople, *J. Chem. Phys.* **106**, 1063 (1997)
- [10] A. D. Becke, *J. Comput. Chem.* **20**, 63 (1999) and references therein.; see also A. D. Becke, *J. Chem. Phys.* **107**, 8554 (1997); H. L. Schmider and A. D. Becke, *J. Chem. Phys.* **108**, 9624 (1998); *ibid.* **109**, 8188 (1998).
- [11] *NIST-JANAF Thermochemical Tables, 4th Edition*, Ed. M. W. Chase Jr., *J. Phys. Chem. Ref. Data* Monograph 9 (1998).

- [12] J.D. Cox, D.D. Wagman, and V.A. Medvedev, *CODATA key values for thermodynamics* (Hemisphere, New York, 1989)
- [13] J. M. L. Martin, Chem. Phys. Lett. **259**, 669 (1996)
- [14] J. M. L. Martin and P. R. Taylor, J. Chem. Phys. **106**, 8620 (1997)
- [15] J. M. L. Martin, Chem. Phys. Lett. **242**, 343 (1995)
- [16] C. Schwartz, in *Methods in Computational Physics 2*, Ed. B. J. Alder (Academic Press, New York, 1963).
- [17] R. N. Hill, J. Chem. Phys. **83**, 1173 (1985)
- [18] W. Kutzelnigg and J. D. Morgan III, J. Chem. Phys. **96**, 4484 (1992); erratum **97**, 8821 (1992)
- [19] J. M. L. Martin and P. R. Taylor, J. Phys. Chem. A **102**, 2995 (1998)
- [20] T. Helgaker, W. Klopper, H. Koch, and J. Noga, J. Chem. Phys. **106**, 9639 (1997)
- [21] C. W. Bauschlicher Jr. and A. Ricca, J. Phys. Chem. A **102**, 8044 (1998)
- [22] T. H. Dunning Jr., "Basis sets: correlation consistent" in *Encyclopedia of Computational Chemistry*, 5 vols. (Wiley, NY, 1998); see also D. Feller and K. A. Peterson, J. Chem. Phys. **108**, 154 (1998)
- [23] J. M. L. Martin, J. Chem. Phys. **108**, 2791 (1998)
- [24] C. W. Bauschlicher Jr. and H. Partridge, Chem. Phys. Lett. **240**, 533 (1995)
- [25] H.-J. Werner, and P. J. Knowles, MOLPRO 97.3, a package of *ab initio* programs, with contributions from J. Almlöf, R. D. Amos, A. Berning, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, S. T. Elbert, C. Hampel, R. Lindh, A. W. Lloyd, W. Meyer, A. Nicklass, K. A. Peterson, R. M. Pitzer, A. J. Stone, P. R. Taylor, M. E. Mura, P. Pulay, M. Schütz, H. Stoll, Thorsteinsson, T.

- [26] MOLPRO is a package of ab initio programs written by H.-J. Werner and P. J. Knowles, with contributions from J. Almlöf, R. D. Amos, A. Berning, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, S. T. Elbert, C. Hampel, R. Lindh, A. W. Lloyd, W. Meyer, A. Nicklass, K. Peterson, R. Pitzer, A. J. Stone, P. R. Taylor, M. E. Mura, P. Pulay, M. Schütz, H. Stoll, and T. Thorsteinsson.
- [27] J. Noga and R. J. Bartlett, *J. Chem. Phys.* **86**, 7041 (1987) erratum **89**, 3401 (1988)
- [28] J. F. Stanton, J. Gauss, J. D. Watts, W. Lauderdale, and R. J. Bartlett, (1996) ACES II, an ab initio program system, incorporating the MOLECULE vectorized molecular integral program by Almlöf, J., and Taylor, P. R., and a modified version of the ABACUS integral derivative package by T. Helgaker, H. J. Aa. Jensen, P. Jørgensen, J. Olsen, and P. R. Taylor, .
- [29] T. H. Dunning Jr., *J. Chem. Phys.* **90**, 1007 (1989)
- [30] D. E. Woon and T. H. Dunning Jr., *J. Chem. Phys.* **98**, 1358 (1993).
- [31] R. A. Kendall, T. H. Dunning, and R. J. Harrison, *J. Chem. Phys.* **96**, 6796 (1992)
- [32] G. D. Purvis III and R. J. Bartlett, *J. Chem. Phys.* **76**, 1910 (1982)
- [33] K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, *Chem. Phys. Lett.* **157**, 479 (1989)
- [34] J. D. Watts, J. Gauss, and R. J. Bartlett, *J. Chem. Phys.* **98**, 8718 (1993)
- [35] T. J. Lee and G. E. Scuseria, in *Quantum mechanical electronic structure calculations with chemical accuracy*, Ed. S. R. Langhoff (Kluwer, Dordrecht, The Netherlands, 1995), pp. 47–108.
- [36] D. E. Woon and T. H. Dunning Jr., *J. Chem. Phys.* **103**, 4572 (1995)
- [37] J. M. L. Martin and P. R. Taylor, *Chem. Phys. Lett.* **225**, 473 (1994)

- [38] R. J. Gdanitz and R. Ahlrichs, *Chem. Phys. Lett.* **143**, 413 (1988)
- [39] R. D. Cowan and M. Griffin, *J. Opt. Soc. Am.* **66**, 1010 (1976)
- [40] R. L. Martin, *J. Phys. Chem.* **87**, 750 (1983)
- [41] M. Douglas and N. M. Kroll, *Ann. Phys. (NY)* **82**, 89 (1974); R. Samzow, B. A. Heß, and G. Jansen, *J. Chem. Phys.* **96**, 1227 (1992) and references therein.
- [42] C. L. Collins and R. S. Grev, *J. Chem. Phys.* **108**, 5465 (1998)
- [43] C. W. Bauschlicher Jr., *Theor. Chem. Acc.* **101**, 421 (1999)
- [44] B. A. Heß, C. M. Marian, and S. D. Peyerimhoff, in *Modern Electronic Structure Theory, Vol. 1*, (Yarkony, D. R., Ed.); World Scientific, Singapore, **1995**, p. 152–278.
- [45] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. G. Johnson, W. Chen, M. W. Wong, J. L. Andres, M. Head-Gordon, E. S. Replogle, and J. A. Pople, *Gaussian 98, Revision A.3* (Gaussian, Inc., Pittsburgh, PA, 1998).
- [46] A. D. Becke, *J. Chem. Phys.* **98**, 5648 (1993).
- [47] C. Lee, W. Yang, and R. G. Parr, *Phys. Rev. B* **37**, 785 (1988)
- [48] J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais, *Phys. Rev. B* **46**, 6671 (1992) and references therein

- [49] J. M. L. Martin and O. Uzan, Chem. Phys. Lett. **282**, 16 (1998)
- [50] J. M. L. Martin, J. Phys. Chem. A **102**, 1394 (1998)
- [51] J. M. L. Martin, Spectrochim. Acta A **55**, 713 (1999)
- [52] J. M. L. Martin and P. R. Taylor, Mol. Phys. **96**, 681 (1999)
- [53] J. M. L. Martin, in NATO ASI Symposium Volume *Energetics of stable molecules and reactive intermediates* (ed. M. E. Minas da Piedade), NATO ASI Series C **535** (Kluwer, Dordrecht, 1999), pp. 373-415.
- [54] F. Jensen, J. Chem. Phys., submitted.
- [55] D. Feller, J. Chem. Phys. **96**, 6104 (1992)
- [56] A. Halkier, T. Helgaker, P. Jørgensen, W. Klopper, H. Koch, J. Olsen, and A. K. Wilson, Chem. Phys. Lett. **286**, 243 (1998)
- [57] R. J. Gdanitz, J. Chem. Phys. **110**, 706 (1999)
- [58] G. de Oliveira, J. M. L. Martin, F. De Proft, and P. Geerlings, Phys. Rev. A, in press.
- [59] M. Scheer, R. C. Bilodeau, and H. K. Haugen, Phys. Rev. Lett. **80**, 2562 (1998)
- [60] D. Calabrese, A.M. Covington, and J.S. Thompson, Phys. Rev. A **54**, 2797 (1996).
- [61] M. Scheer, R.C. Bilodeau, J. Thogersen, and H.K. Haugen, Phys. Rev. A **57**, R1493 (1998).
- [62] J. Thogersen, L.D. Steele, M. Scheer, C.A. Brodie, and H.K. Haugen, J. of Phys. B, **29**, 1323 (1996).
- [63] K. A. Peterson, D. E. Woon, and T. H. Dunning Jr., J. Chem. Phys. **100**, 7410 (1994).
- [64] C. E. Moore, *Atomic energy levels*, Natl. Bur. Stand. (US) Circ. **1949**, 467.
- [65] R. D. Johnson III and J. F. Liebman, personal communication.

- [66] T. J. Lee and P. R. Taylor, *Int. J. Quantum Chem. Symp.* **23**, 199 (1989)
- [67] J. M. L. Martin, J. El-Yazal, and J. P. François, *J. Phys. Chem.* **100**, 15358 (1996)
- [68] A. P. Scott and L. Radom, *J. Phys. Chem.* **100**, 16502 (1996)
- [69] W. Klopper, J. Noga, H. Koch, and T. Helgaker, *Theor. Chem. Acc.* **97**, 164 (1997)
- [70] e.g. H. Koch, A. Sanchez de Meras, T. Helgaker, and O. Christiansen, *J. Chem. Phys.* **104**, 4157 (1996)
- [71] K. P. Huber and G. Herzberg, *Constants of Diatomic Molecules* (Van Nostrand Reinhold, New York, 1979)
- [72] P. Geerlings, F. De Proft, and J. M. L. Martin, in *Theoretical and Computational Chemistry, Vol. 4: Recent developments and applications of modern density functional theory* (ed. J. Seminario), Elsevier, New York, 1996, pp. 773.
- [73] D. G. Truhlar, *Chem. Phys. Lett.* **294**, 45 (1998)
- [74] J. M. L. Martin, J. El-Yazal, and J. P. François, *Mol. Phys.* **86**, 1437 (1995).
- [75] H.Y. Afeefy, J.F. Liebman, and S.E. Stein, "Neutral Thermochemical Data" in NIST Chemistry WebBook, NIST Standard Reference Database Number 69, Eds. W.G. Mallard and P.J. Linstrom, November 1998, National Institute of Standards and Technology, Gaithersburg MD, 20899 (<http://webbook.nist.gov>).

**Table I.** Computed and observed electron affinities (eV) of first and second-row atoms.

	SCF limit A+B.C <sup>n</sup>	CCSD(T) valence correlation limit A+B/I <sup>3</sup> [a]	Inner-shell correlation CCSD(T)/MTav5z	Spin-orbit coupling MRCI/Mtav5z	Scalar relativ. effects ACPF/MTavqz	FCI correction for CCSD(T) imperfections	Best calculated electron affinity (This work)	Experimental elec. affin. [b]
H	-0.32877	1.08297	0.00000	0.00000	-0.00004	0.00000	0.75416	0.7542(2)
B	-0.26754	0.52465	0.00427	-0.00060	-0.00127	0.01907	0.27858	0.27972(3) [c]
C	0.54826	0.70047	0.00720	-0.00332	-0.00283	0.01309	1.26288	1.2629(3)
N	--	--	--	--	--	--	--	--
O	-0.53902	1.99391	0.00173	-0.00222	-0.00588	0.01223	1.46075	1.461122(3)
F	1.30727	2.11864	0.00430	-0.01652	-0.00928	0.00056	3.40496	3.401190(4)
Al	0.04101	0.40219	-0.01617	-0.00385	-0.00536	0.01497	0.43277	0.43283(5) [d]
Si	0.95579	0.46046	-0.00965	-0.01806	-0.00787	0.00992	1.39060	1.38946(6) [e]
P	-0.45732	1.19312	-0.00521	0.01229	-0.00937	0.01124	0.74474	0.7465(3)
S	0.90392	1.18504	-0.00161	-0.00410	-0.01223	0.00441	2.07544	2.077104(1)
Cl	2.52998	1.13504	0.00085	-0.03657	-0.01509	-0.00309	3.61113	3.61269(6)
Recently calculated values dating back to 1995								
	G3 theory	Ref. f	Ref. g	Ref. h	Ref. i	Ref. j	Ref. k	Ref. l
H		0.7542	0.747		0.637			
B	0.204	0.2833	0.241	0.2038	0.282	0.260	0.279	0.2795
C	1.193	1.2655	1.259	1.1925	1.220	1.210		
N								
O	1.336	1.4540	1.432	1.3356	1.292			
F	3.400	3.3980	3.395	3.3997	2.180			
Al	0.390		0.433	0.3903	0.450	0.433	0.427	
Si	1.379		1.405	1.3790	1.372			
P	0.711		0.714	0.7112	0.748	0.702		
S	2.064		2.059	2.0641	1.996			
Cl	3.608		3.623	3.6079	3.332			

a) For the largest atoms, P, S, and Cl, the bases were doubly augmented, i.e. the daug-ccpVnZ rather than the aug-ccpVnZ series was used.

b) Unless otherwise indicated, experimental values are those from CRC Handbook of Chemistry and Physics, 78th Edition (CRC Press, Boca Raton, FL, 1997).

c) M. Scheer, R. C. Bilodeau, and H K Haugen, *Phys. Rev. Lett.* **80**, 2562 (1998).

d) M. Scheer, R.C. Bilodeau, J. Thogersen, and H.K. Haugen, *Phys. Rev. A* **57**, 1493 (1998).

e) J. Thogersen, L.D. Steele, M. Scheer, C.A. Brodie, and H.K. Haugen, *J. of Phys. B* **29**, 1323 (1996).

f) R. J. Gdanitz, *J. Chem. Phys.* **110**, 706 (1999).

g) G. L. Gutsev, P. Jena, and R. J. Bartlett, *Chem. Phys. Lett.* **291**, 547 (1998).

h) C. Guo-xin, P. P. Ong, and L. Ting, *Chem. Phys. Lett.* **290**, 211 (1998).

i) W. P. Wijesundera and F. A. Parpia, *Phys. Rev. A* **57**, 3462 (1998).

j) W. P. Wijesundera, *Phys. Rev. A* **55**, 1785 (1997).

k) E. Eliav, Y. Ishikawa, P. Pyykkö, and U. Kaldor, *Phys. Rev. A* **56**, 4532 (1997).

l) C. Froese Fischer, A. Ynnerman, and G. Gaigalas, *Phys. Rev. A* **51**, 4611 (1995).

**Table II.** Complete basis set limit extrapolations for valence correlation contributions to electron affinities (eV) [a].

Halkier et al. A+B/I <sup>3</sup>				Martin A+B/I <sup>c</sup>			Feller A+B/C <sup>i</sup>		
	TQ	Q5	56	DTQ	TQ5	Q56	DTQ	TQ5	Q56
H	1.07834	1.08189	1.08297	1.07745	1.08778	1.08458	1.07382	1.08188	1.08214
B	0.52100	0.52406	0.52465	0.51709	0.52518	0.52465	0.51520	0.52323	0.52339
C	0.70022	0.70042	0.70047	0.69644	0.69997	0.70009	0.69202	0.69721	0.69879
N	---	---	---	---	---	---	---	---	---
O	1.99387	1.99306	1.99391	2.03687	1.99034	1.99321	1.99095	1.97966	1.98747
F	2.12556	2.11895	2.11864	2.20763	2.11058	2.11731	2.15322	2.10414	2.11207
Al	0.40385	0.40219	---	0.39990	0.39885	---	0.39797	0.39997	---
Si	0.46248	0.46046	---	0.45734	0.45672	---	0.45269	0.45675	---
P	1.20355	1.19166	---	1.20887	1.17356	---	1.18412	1.17749	---
S	1.19947	1.18400	---	1.24532	1.16889	---	1.21649	1.16716	---
Cl	1.15231	1.13398	---	1.27950	1.11802	---	1.66795	1.11438	---
Mean. Sgn. Error [b]	0.0042	-0.0008	0.0004	0.0328	-0.0068	0.0003	0.0546	-0.0096	-0.0029
Mean. Abs. Error [b]	0.0063	0.0018	0.0011	0.0373	0.0078	0.0013	0.0653	0.0096	0.0029
No extrapolation			A+B/(I+1/2) <sup>4</sup> +C/(I+1/2) <sup>6</sup> [c]				A + B exp[-(I-1)] + C exp[-(I-1) <sup>2</sup> ] [d]		
	Q	5	6	DTQ	TQ5	Q56	DTQ	TQ5	Q56
H	1.07023	1.07592	1.07889	1.07731	1.08261	1.08305	1.07617	1.07925	1.08062
B	0.51191	0.51784	0.52071	0.51873	0.52456	0.52448	0.51826	0.52131	0.52238
C	0.68680	0.69344	0.69641	0.69749	0.69967	0.70002	0.69634	0.69731	0.69813
N	---	---	---	---	---	---	---	---	---
O	1.93908	1.96543	1.97743	2.00302	1.98945	1.99239	1.98349	1.98074	1.98441
F	2.06733	2.09252	2.10353	2.14066	2.11313	2.11669	2.11597	2.10714	2.10993
Al	0.39513	0.39857	---	0.40100	0.40108	---	0.40105	0.40057	---
Si	0.44791	0.45404	---	0.45780	0.45888	---	0.45782	0.45759	---
P	1.14764	1.16912	---	1.20637	1.18416	---	1.19121	1.18155	---
S	1.13267	1.15772	---	1.21485	1.17461	---	1.18794	1.17221	---
Cl	1.07448	1.10352	---	1.18290	1.12293	---	1.14227	1.12030	---
Mean. Sgn. Error [b]	-0.0325	-0.0170	-0.0083	0.0102	-0.0047	-0.0004	-0.0028	-0.0080	-0.0046
Mean. Abs. Error [b]	0.0325	0.0170	0.0083	0.0139	0.0047	0.0011	0.0044	0.0080	0.0046

a) Here only the aug-cc-pVnZ series was used.

b) Relative to Expt.-all contributions other than valence correlation (from Table I)

c) Originally proposed by J. M. L. Martin, *Chem. Phys. Lett.* **259**, 669 (1996)d) Originally proposed by K. A. Peterson, D. E. Woon, and T. H. Dunning, Jr., *J. Chem. Phys.* **100**, 7410 (1994)



**Table III.** Calculated and experimental spin-orbit contributions to EA (eV).

	Calc. spin-orbit splitting (CASCI/MTav5z) [a]	Spin-orbit splitting from experimental fine structure
H	--	--
B	-0.00060	-0.00055
C	-0.00332	-0.00367
N	--	--
O	-0.00222	-0.00235
F	-0.01652	-0.01670
Al	-0.00385	-0.00361
Si	-0.01806	-0.01856
P	0.01229	0.01110
S	-0.00410	-0.00429
Cl	-0.03657	-0.03647

a) For technical reasons, only s, p, d, and f orbitals were used.

**Table IV.** Calculated (W[Q5;Q5;TQ5] level) and experimental total atomization energies (kcal/mol).

	SCF limit (TQ5 extr.)	CCSD(T) val. Corr. limit (Q5 extr.)	Inner-shell correlation CCSD(T) [c]	Expmatl. spin-orbit splitting	Scalar relativ. effects ACPF [d]	Best calc. Total Atomiz. Energy (TAE)	Error in calculated TAE's	Experimental TAE <sub>e</sub> [e]	Experimental error	Zero-Point Vibrational Energies [m]	Experimental TAE <sub>0</sub>
H <sub>2</sub>	83.85	25.68	0.00	0.00	0.00	109.53	0.05	109.48	0.00 [f]	6.21	103.27
N <sub>2</sub>	119.67	107.72	0.75	0.00	-0.12	228.03	-0.39	228.42	0.04 [f]	3.36	225.06
O <sub>2</sub>	26.76	93.20	0.24	-0.45	-0.16	119.60	-0.62	120.22	0.04 [f]	2.25	117.97
F <sub>2</sub>	-31.09	69.55	-0.09	-0.77	0.03	37.63	-0.61	38.24	0.10 [f]	1.30	36.94
HF	100.03	41.52	0.18	-0.39	-0.20	141.15	-0.03	141.18	0.17 [f]	5.85	135.33
CH	57.22	26.73	0.14	-0.04	-0.04	84.01	0.07	83.94	0.23 [f]	4.04	79.90
CO	181.84	76.89	0.94	-0.31	-0.14	259.21	-0.06	259.27	0.12 [g]	3.11	256.16
NO	54.90	97.01	0.40	-0.05	-0.17	152.09	-0.44	152.54	0.03 [g]	2.71	149.82
CS	104.15	67.01	0.75	-0.64	-0.16	171.11	-0.33	171.45	0.23 [f]	1.83	169.41
SO	53.12	72.84	0.46	-0.78	-0.32	125.32	0.09	125.22	0.04 [f]	1.64	123.58
HCl	80.86	26.62	0.20	-0.84	-0.26	106.57	0.09	106.48	0.02 [f]	4.24	102.24
ClF	15.38	47.29	0.08	-1.23	-0.14	61.39	-0.08	61.48	0.01 [f]	1.12	60.36
Cl <sub>2</sub>	26.84	33.05	0.19	-1.68	-0.18	58.22	0.25	57.98	0.00 [f]	0.80	57.18
HNO	85.42	119.67	0.40	-0.22	-0.25	205.03	-0.38	205.41	0.06 [h]	8.56	196.85
CO <sub>2</sub>	258.05	130.29	1.64	-0.53	-0.46	389.00	-0.15	389.15	0.06 [g]	7.24	381.91
H <sub>2</sub> O	160.01	72.73	0.37	-0.22	-0.26	232.62	0.02	232.60	0.12 [g]	13.25	219.35
H <sub>2</sub> S	133.64	50.03	0.34	-0.56	-0.41	183.04	0.44	182.60	0.12 [i]	9.40 [n]	173.15
HOCl	86.67	79.38	0.31	-1.06	-0.30	165.00	0.18	164.81	0.12 [i]	8.21 [o]	156.61
OCS	218.24	116.09	1.68	-0.87	-0.54	334.60	0.36	334.24	0.48 [i]	5.72 [k]	328.53
ClCN	169.49	114.31	1.76	-0.93	-0.44	284.20	-0.32	284.52	0.48 [i]	5.33 [p]	279.20
SO <sub>2</sub> [a]	122.06	137.61	0.67	-1.01	-0.71	258.62	0.23	258.39	0.08 [i]	4.38 [a]	253.92
CH <sub>3</sub>	243.43	63.42	1.04	-0.08	-0.17	307.63	0.19	307.44	0.10 [j]	18.44 [q]	289.00
NH <sub>3</sub>	203.28	94.19	0.62	0.00	-0.25	297.84	-0.22	298.06	0.13 [g]	21.33	276.73
PH <sub>3</sub>	173.23	68.63	0.30	0.00	-0.47	241.69	0.13	241.56	0.41 [k]	14.44 [r]	227.13
C <sub>2</sub> H <sub>2</sub>	299.88	103.15	2.44	-0.17	-0.27	405.02	-0.34	405.36	0.24 [l]	16.46	388.90
H <sub>2</sub> CO	264.82	108.53	1.25	-0.31	-0.32	373.98	0.15	373.82	0.12 [i]	16.53	357.25
CH <sub>4</sub>	331.58	87.67	1.21	-0.08	-0.19	420.20	0.09	420.11	0.14 [k]	27.60	392.51
C <sub>2</sub> H <sub>4</sub> [b]	435.07	126.84	2.36	-0.17	-0.33	563.77	0.26	563.51	0.17 [k]	31.60	531.91

a) The geometry and zero point energy are from J. M. L. Martin, *J. Chem. Phys.* **108** (1998) 2791.

b) The geometry and zero point energy are from J. M. L. Martin and P. R. Taylor, *Chem. Phys. Lett.* **248** (1996) 336. Valence energies were obtained with VnZ set.

c) The basis sets were acv5z for most first-row molecules and MTavqz for most second-row.

For SO<sub>2</sub>, CH<sub>3</sub>, NH<sub>3</sub>, C<sub>2</sub>H<sub>2</sub>, H<sub>2</sub>CO, CH<sub>4</sub> and C<sub>2</sub>H<sub>4</sub>, the MT basis set was used.

d) The basis sets were acvqz for most first-row molecules and MTavqz for most second-row.

For SO<sub>2</sub>, CH<sub>3</sub>, NH<sub>3</sub>, C<sub>2</sub>H<sub>2</sub>, H<sub>2</sub>CO, CH<sub>4</sub> and C<sub>2</sub>H<sub>4</sub>, the MT basis set was used.

- e) This "bottom of the well" value was obtained by adding the experimental TAE\_0 to the best available zero point vibration energy.
- f) K. P. Huber and G. Herzberg, "Constants of Diatomic Molecules" (Van Nostrand Reinhold, New York, 1979).
- g) J. D. Cox, D. D. Wagman, and V. A. Medvedev, CODATA *key values for thermodynamics* (Hemisphere, New York, 1989).
- h) R. N. Dixon, J. Chem. Phys. **104**, 6905 (1996).
- i) L. V. Gurvich, I. V. Veyts, and C. B. Alcock (Eds.), Thermodynamic properties of individual substances, 4th Ed. (Hemisphere, New York, 1989).
- j) J. M. L. Martin, Chem. Phys. Lett. **273**, 98 (1997).
- k) NIST-JANAF *Thermochemical Tables, 4th Edition*, Ed. M. W. Chase Jr., J. Phys. Chem. Ref. Data Monograph 9 (1998).
- l) D. D. Wagman, W. H. Evans, V. B. Parker, R. H. Schumm, I. Halow, S. M. Bailey, K. L. Churney, and R. L. Nuttall, J. Phys. Chem. Ref. Data **11**, supplement 2 (1982).
- m) ZPVE for diatomics were derived from spectroscopic constants in [f], while those for polyatomics were taken from [h]
- n) I. Kozin and P. Jensen, J. Mol. Spec. **163** (1994) 483.
- o) B. Abel, H.H. Hamann, A.A. Kachanov, and J. Troe, J. Chem. Phys. **104** (1996) 3189.
- p) T.J. Lee, J.M.L. Martin, C.E. Dateo and P. R. Taylor, J. Phys Chem. **99** (1995) 15858.
- q) Taken from unpublished work of D. W. Schwenke, A. Pradhan, and H. Partridge, quoted in C.W. Bauschlicher and H. Partridge, J. Chem. Phys. **103** (1995) 10589.
- r) Quartic force field calculated at CCSD(T)/VQZ+1 level for the purposes of this work.

**Table V.** Comparison between the W[Q5;Q5;TQ5] and W[56;56;Q56] approaches [a]. Energies are in kcal/mol.

	W[56;56;Q56] Tot. Atom. Energy [a]	Error	W[Q5;Q5;TQ5] Tot. Atom. Energy [a]	Error	W[56;56;Q56] – W[Q5;Q5;TQ5]
N <sub>2</sub>	227.94	-0.48	228.03	-0.39	-0.09
O <sub>2</sub>	119.68	-0.54	119.60	-0.62	0.08
F <sub>2</sub>	37.67	-0.57	37.63	-0.61	0.04
HF	141.14	-0.04	141.15	-0.03	0.00
CO	259.18	-0.09	259.21	-0.06	-0.03
NO	152.12	-0.42	152.09	-0.44	0.02
H <sub>2</sub> O	232.55	-0.05	232.62	0.02	-0.08

a) See text for definition of approaches.

**Table VI.** Comparison of computational approaches for the evaluation of TAE (kcal/mol).

	Experimental TAE <sub>e</sub>	Best W[Q5;Q5;TQ5]	3 pt. extr. W[TQ5;TQ5;TQ5]	W2 W[TQ,Q5,TQ5]	MT W[TQ,TQ, DTQ]3.182	MT W[DT,TQ, DTQ]3.26	W1 W[DT,TQ, DTQ]3.22	G3	G2	CBS-QB3
H <sub>2</sub>	109.48	109.53	109.47	109.53	109.53	109.52	109.53	109.71	110.32	110.70
N <sub>2</sub>	228.42	228.03	227.90	228.06	227.87	227.88	227.89	226.50	227.32	227.72
O <sub>2</sub>	120.22	119.60	119.43	119.58	119.79	119.83	119.89	119.42	118.11	120.99
F <sub>2</sub>	38.24	37.63	37.51	37.64	37.70	37.74	37.72	37.84	38.19	39.10
HF	141.18	141.15	140.95	141.16	141.48	141.53	141.59	141.05	141.85	141.99
CH	83.94	84.01	83.90	84.02	84.07	84.05	84.14	85.22	84.39	83.92
OO	259.27	259.21	259.05	259.15	259.27	259.27	259.35	259.54	261.11	259.94
NO	152.54	152.09	151.95	152.06	152.21	152.22	152.28	152.55	153.40	153.73
CS	171.45	171.11	171.57	170.94	170.57	170.49	170.53	172.41	172.33	172.16
SO	125.22	125.32	125.69	125.24	124.63	124.55	124.60	124.61	122.48	126.28
HCl	106.48	106.57	106.47	106.52	106.55	106.53	106.53	106.17	106.68	107.45
ClF	61.48	61.39	61.42	61.39	61.30	61.27	61.24	60.72	62.17	61.61
Cl <sub>2</sub>	57.98	58.22	58.67	58.18	57.54	57.31	57.25	56.83	56.56	59.58
HNO	205.41	205.03	204.73	205.03	205.21	205.21	205.20	204.93	206.91	206.51
CO <sub>2</sub>	389.15	389.00	388.59	389.01	389.27	389.27	389.43	390.30	391.76	391.15
H <sub>2</sub> O	232.60	232.62	232.30	232.64	233.00	233.01	233.04	231.91	232.51	233.08
H <sub>2</sub> S	182.60	183.04	182.86	182.92	183.01	182.92	182.92	181.95	182.27	183.52
HOCl	164.81	165.00	164.81	164.98	164.92	164.84	164.86	164.01	164.93	166.06
OCS	334.24	334.60	334.57	334.44	334.18	334.09	334.17	336.11	336.03	337.06
ClCN	284.52	284.20	284.14	284.12	283.75	283.64	283.72	284.76	285.19	285.94
SO <sub>2</sub>	258.39	258.62	260.20	258.61	257.09	257.07	257.38	254.66	253.43	259.01
CH <sub>3</sub>	307.44	307.63	307.35	307.65	307.81	307.74	307.82	307.62	306.49	307.00
NH <sub>3</sub>	298.06	297.84	297.49	297.93	298.15	298.09	298.10	296.64	297.26	297.48
PH <sub>3</sub>	241.56	241.69	241.38	241.58	241.82	241.65	241.66	240.00	241.08	242.64
C <sub>2</sub> H <sub>2</sub>	405.36	405.02	404.54	404.94	405.19	405.13	405.23	404.63	403.73	403.90
H <sub>2</sub> CO	373.82	373.98	373.66	374.05	374.21	374.17	374.06	374.27	375.63	375.05
CH <sub>4</sub>	420.11	420.20	419.85	420.22	420.34	420.27	420.37	419.53	419.94	420.11
C <sub>2</sub> H <sub>4</sub>	563.51	563.77	563.23	563.70	563.78	563.63	563.51	562.80	562.37	562.81
mean abs. error		0.235	0.399	0.228	0.344	0.340	0.330	0.861	1.175	0.906
mean abs. Error (w/out SO <sub>2</sub> )		0.236	0.346	0.228	0.308	0.303	0.305	0.755	1.035	0.917
(a)		0.185	0.304	0.179	0.288	0.289	0.296	0.804	1.069	0.914
Mean abs. Error (1st row)		0.240	0.402	0.247	0.268	0.245	0.255	0.674	1.095	0.838
without F <sub>2</sub> , O <sub>2</sub> , and NO		0.172	0.339	0.177	0.233	0.211	0.231	0.731	1.114	0.815
Mean abs. Error (2nd row)		0.228	0.393	0.198	0.461	0.486	0.445	1.151	1.298	1.013
without H <sub>2</sub> S and SO <sub>2</sub>		0.205	0.250	0.182	0.373	0.411	0.397	0.919	0.999	1.067
Max. abs. Error		0.623	1.812	0.641	1.304	1.319	1.012	3.736	4.960	2.816
(other than SO <sub>2</sub> )		0.623	0.787	0.641	0.776	0.957	0.799	1.923	2.740	2.816

(a) without F<sub>2</sub>, O<sub>2</sub>, NO, H<sub>2</sub>S, and SO<sub>2</sub>

**Table VII.** Comparison of basis sets and correlation methods for core-correlation contributions to TAE (kcal/mol).

	CCSD(T)/ largest bases [a]	CCSD(T)/ MT	CCSD(T)/ MTsmall	CCSD(T) G3large	CCSD/ MT
H <sub>2</sub>	0.00	0.00	0.00	0.00	0.00
N <sub>2</sub>	0.75	0.74	0.82	1.22	0.44
O <sub>2</sub>	0.24	0.24	0.24	0.59	0.03
F <sub>2</sub>	-0.09	-0.08	-0.08	0.11	-0.34
HF	0.18	0.17	0.18	0.23	0.14
CH	0.14	0.14	0.14	0.23	0.09
CO	0.94	0.92	0.90	1.23	0.72
NO	0.40	0.41	0.41	0.89	0.14
CS	0.75	0.70	0.66	1.23	0.23
SO	0.46	0.39	0.42	0.91	0.14
HCl	0.20	0.16	0.15	0.36	0.09
ClF	0.08	0.05	0.09	0.42	-0.22
Cl <sub>2</sub>	0.19	0.16	0.18	0.69	-0.24
HNO	0.40	0.40	0.41	0.91	0.10
CO <sub>2</sub>	1.64	1.68	1.67	2.52	1.38
H <sub>2</sub> O	0.37	0.36	0.37	0.46	0.30
H <sub>2</sub> S	0.34	0.27	0.25	0.56	0.14
HOCl	0.31	0.26	0.29	0.69	-0.04
OCS	1.68	1.60	1.58	2.60	1.18
ClCN	1.76	1.70	1.71	2.46	1.29
SO <sub>2</sub>	0.67	0.67	0.78	1.81	0.10
CH <sub>3</sub>	1.04	1.04	1.04	1.10	1.04
NH <sub>3</sub>	0.62	0.62	0.64	0.82	0.53
PH <sub>3</sub>	0.30	0.23	0.22	0.58	0.08
C <sub>2</sub> H <sub>2</sub>	2.44	2.36	2.34	2.41	2.19
H <sub>2</sub> CO	1.25	1.25	1.26	1.62	1.06
CH <sub>4</sub>	1.21	1.21	1.21	1.80	1.20
C <sub>2</sub> H <sub>4</sub>	2.36	2.36	2.27	2.45	--

a) The MTavqz basis set was used for most second-row, and acv5z for most first-row molecules.

For SO<sub>2</sub>, CH<sub>3</sub>, NH<sub>3</sub>, C<sub>2</sub>H<sub>2</sub>, H<sub>2</sub>CO, CH<sub>4</sub> and C<sub>2</sub>H<sub>4</sub>, the MT basis set was used.

**Table VIII.** Comparison of Zero-Point Vibrational Energies (kcal/mol)

	Best available ZPVE [a]	B3LYP/cc-pVTZ ZPVE	B3LYP/cc-pVTZ ZPVE Scaled by 0.985	Error in scaled ZPVE
H <sub>2</sub>	6.21	6.32	6.22	0.01
N <sub>2</sub>	3.36	3.50	3.45	0.09
O <sub>2</sub>	2.25	2.33	2.29	0.04
F <sub>2</sub>	1.30	1.50	1.48	0.18
HF	5.85	5.84	5.76	-0.09
CH	4.04	4.03	3.97	-0.07
CO	3.11	3.16	3.11	0.00
NO	2.71	2.83	2.78	0.07
CS	1.83	1.86	1.84	0.00
SO	1.64	1.64	1.62	-0.02
HCl	4.24	4.20	4.14	-0.10
ClF	1.12	1.12	1.11	-0.01
Cl <sub>2</sub>	0.80	0.77	0.76	-0.04
HNO	8.56	8.68	8.55	-0.01
CO <sub>2</sub>	7.24	7.34	7.23	-0.01
H <sub>2</sub> O	13.25	13.35	13.15	-0.10
H <sub>2</sub> S	9.40 [b]	9.42	9.28	-0.11
HOCl	8.21 [c]	8.26	8.13	-0.07
OCS	5.72 [d]	5.78	5.70	-0.02
ClCN	5.33 [e]	5.52	5.44	0.11
SO <sub>2</sub>	4.38 [f]	4.42	4.36	-0.02
CH <sub>3</sub>	18.44 [g]	18.61	18.33	-0.11
NH <sub>3</sub>	21.33	21.49	21.17	-0.16
PH <sub>3</sub>	14.44 [h]	14.96	14.73	0.30
C <sub>2</sub> H <sub>2</sub>	16.46	16.94	16.68	0.22
H <sub>2</sub> CO	16.53	16.65	16.40	-0.13
CH <sub>4</sub>	27.60	28.00	27.58	-0.02
C <sub>2</sub> H <sub>4</sub>	31.60	31.96	31.48	-0.12

a) ZPVE for diatomics were derived from spectroscopic constants in Hub.79, while those for polyatomics were taken from J. M. L. Martin, *J. Molec. Struct. (THEOCHEM)* 398-399, 135-144 (1997) [WATOC'96 special issue]., unless otherwise indicated.

b) I. Kozin and P. Jensen, *J. Mol. Spec.* **163** (1994) 483.

c) B. Abel, H.H. Hamann, A. A. Kachanov, and J. Troe, *J. Chem. Phys.* **104** (1996) 3189.

d) Derived from spectroscopic constants in NIST-JANAF *Thermochemical Tables, 4th Edition*, Ed. M. W. Chase Jr., J. Phys. Chem. Ref. Data Monograph 9 (1998).

e) T. J. Lee, J. M. L. Martin, C. E. Dateo and P. R. Taylor, *J. Phys Chem.* **99** (1995) 15858.

f) J. M. L. Martin, *J. Chem. Phys.* **108** (1998) 2791.

g) Taken from unpublished work of D. W. Schwenke, A. D. Pradhan, and H. Partridge, quoted in C.W. Bauschlicher and H. Partridge, *J. Chem. Phys.* **103** (1995) 10589.

h) Quartic force field calculated at CCSD(T)/VQZ+1 level for the purposes of this work.

**Table IX.** Examples of W1 calculations for some larger organic molecules. All energies are in kcal/mol.

	Benzene	Trans-butadiene
SCF limit	1044.95	775.26
Valence corr. limit	317.62	233.39
core corr.	7.09	4.65
scalar relativistics	-0.97	-0.63
spin-orbit	-0.51	-0.34
Calculated TAE <sub>e</sub>	1368.19	1012.33
ZPVE	62.04	52.51
Calculated TAE <sub>0</sub>	1306.15	959.81
Error	0.04	-0.31
Experimental TAE <sub>0</sub> [d]	1306.11±0.12 [a]	960.13±0.19 [b]
		959.39±0.23 [c]

- a) E. J. Prosen, R. Gilmont, and F. D. Rossini, *J. Res. NBS* **34**, 65 (1945).  
b) E. J. Prosen, F. W. Maron, and F. D. Rossini, *J. Res. NBS* **46**, 106 (1951). Including the error for the heat of formation of carbon brings the total error to ±0.47 kcal/mol.  
c) E. J. Prosen, and F. D. Rossini, *J. Res. NBS* **34**, 59 (1945). Including the error for the heat of formation of carbon brings the total error to ±0.49 kcal/mol.  
d) Experimental data were converted from 298 K to 0 K using H<sub>298</sub>-H<sub>0</sub> data for C<sub>(g)</sub>, C<sub>(gr)</sub>, H<sub>(g)</sub>, and H<sub>2(g)</sub> from reference 12, and computed rigid rotor-harmonic oscillator H<sub>298</sub>-H<sub>0</sub> for the molecule, using the unscaled B3LYP/VTZ harmonic frequencies